

LOAD DISTRIBUTING METHOD, NODE, AND CONTROL PROGRAM**TECHNICAL FIELD**

[0001] The present invention relates to a
5 communication system between two nodes usable plural
communication paths. Particularly, the present invention
relates to a technique of distributing the load between
paths by a transmission node.

BACKGROUND OF THE RELATED ART

[0002] In plural paths between two nodes, the
transmission side dispersively inputs traffics to
respective paths and the reception side reintegrates them
from the respective paths. Thus, a higher transfer rate
15 can be obtained in principle, compared with the case where
a single path is used. This mechanism relates to a
technique of combining paths to construct one logical path.
Typically, Inverse Multiplexing is cited, which is
utilized to obtain a logical broad band line from plural
20 low rate links. As applications to data communications,
there are multiplexing of ATM line and Multilink PPP being
a multiplex protocol for PPP links.

[0003] As a traffic distribution scheme, typified by
Inverse Multiplexing, of forming one logical path by
25 combining paths, the round robin scheme, for example, of

transmitting packets or fragments in turn to respective paths has been broadly used (refer to non-Patent document 1). This method is effective when the rate (or speed) in each path is nearly the same and constant. However, the
5 problem is that when the rates between paths are in imbalance, the transfer rate between nodes is rate-controlled to the slowest path.

[0004] In order to deal with such a problem, the weighted round robin is known for controlling the use
10 frequency of each path according to the rate. The round robin is installed to router products corresponding to the technique of combining paths to form one logical path. The product manufactured by Cisco employs the method of implementing optimum load distribution by hashing input
15 packets to determine the output path. These sets of information are described in the documents published by companies (for example, non-Patent document 2).

[0005] Any one of the above-mentioned well-known techniques assumes that each path consists of a cable link
20 only.

[0006] As to the technique, which combines paths to construct one logical path, the path containing a radio link is proposed. In this proposed scheme, the input packet is divided into fragments to allocate them to each
25 path but the ratio of each fragment size is changed

according to only the rate of each path in transmission
(for example, non-Patent document 3 and Patent document 6).

[0007] In the plural path selectable communications
system, the technique has been proposed of monitoring the
5 status of each path and selecting a path, in consideration
of the path status or the delay time. (For example, refer
to Patent document 1, Patent document 2, Patent document 3,
Patent document 4, and Patent document 5)

Patent document 1: Japanese Patent laid-open
10 publication No. 2000-49862

Patent document 2: Japanese Patent laid-open
publication No. 2001-308917

Patent document 3: Japanese Patent laid-open
publication No. 2001-333100

15 Patent document 4: Japanese Patent laid-open
publication No. 2002-176441

Patent document 5: Japanese Patent laid-open
publication No. 58-27449

Patent document 6: Japanese Patent laid-open
20 publication No. 2000-216815

Non-Patent document 1: "Stripping Within the Network
Subsystem", IEEE Network, July/August, 1995

Non-Patent document 2: Cisco Systems "Load Balancing
with Cisco Express Forwarding", Cisco Application Note,
25 Jan. 1998

Non-patent document 3: "Adaptive Inverse Multiplexing
for Wide-Area Wireless Networks", Snoeren et al,
Proceedings of IEEE GlobeCom, Dec. 1999

5 **DISCLOSURE OF THE INVENTION**

PROBLEMS TO BE SOLVED BY THE INVENTION

10 [0008] In the conventional technique in which radio
links are in the paths between two nodes, the transmission
node refers to the speed or delay of each path as a
selection criterion for transmission path selection.
However, because the path including a radio link
dynamically changing the rate or delay tends to easily
accumulate measurement errors of path status, it is
difficult to accurately predict the time when a packet to
15 be transmitted arrives at the reception side or a
reception completion time. As a result, when the path
selection is carried out according to the prediction
result, a path of which the rate is remarkably decreased
is actually selected. Therefore, the problem is that the
20 multiplexing efficiency decreases.

 [0009] A conventional technique of predicting an
arrival time or a reception completion time of a
transmission packet will be explained below by referring
to Fig. 8.

25 [0010] Referring to Fig. 8, numerals 700-1, 700-2 and

700-3 represent data packets, respectively. A packet transmission history (in solid line) in the transmission node, a packet reception history (in solid line) in the reception side, and prediction (in dotted line) of a packet arrival time and a reception completion time are represented on the time axes.

[0011] First, for actual packet transmission and reception, the transmission node starts transmitting data packet 700-1 at time T1 and completes its transmission operation at time T2. The reception node starts receiving the packet 700-1 at time T3 and completes its reception at time T5.

[0012] Thereafter, the transmission node starts transmitting the data packet 700-2 at time T4 and completes its transmission at time T7. The reception node starts receiving the data packet 700-2 at time T6 because of a decrease of the communication rate and completes its reception at time T11.

[0013] Moreover, the transmission node starts transmitting the data packet 700-3 at time T8 and completes its transmission at time T9. The reception node starts receiving the data packet 700-3 at time T12 and completes its reception at time T13.

[0014] In contrast, in the conventional technique, the packet arrival time and the reception completion time

are predicted based on path status information, such as transmission rate, issued from the reception node, at the reception completion of the data packet 700-1.

[0015] Referring to Fig. 8, the transmission node,
5 for example, receives path status information from the reception node at time TX. In order to predict an arrival time and a reception completion time of the data packet 700-3, the transmission node uses path status information received at time TX. As a result, the transmission node
10 can know the transmission delay I1 and predicts an arrival time of the data packet 700-3 at the time T10 of the transmission time T8 plus the transmission delay I1 of the data packet 700-3.

[0016] However, the actual arrival time of the data
15 packet 700-3 is time T12 and has a margin time of error by ID to the predicted arrival time T10.

[0017] The prediction with a large difference to an actual time causes erroneous path selection, thus resulting in a decrease in multiplexing efficiency.

20 [0018] Such a problem may be solved if the path status at a prediction time can be obtained in real time and is reflected on the transmission time. However, it is impossible actually to receive and transmit a current path status in real time.

25 [0019] The present invention is made in view of the

above-mentioned problems. An object of the present invention is to provide a technique of reflecting a path status obtained from the receiver side on data transmitted in past and predicting an actual data arrival time or an arrival time close to a reception completion time or a reception completion time.

[0020] Another object of the present invention is to provide a technique offering an excellent multiplexing efficiency. In this technique, when data is transferred between two nodes which can select plural paths, the path status obtained from a receiver side is reflected on data transmitted in past. An actual data arrival time or an arrival time close to a reception completion time or a reception completion time is predicted and path selection is performed based on the prediction.

[0021] Further another object of the present invention is to provide a technique of capable of using effective path resources when the performance of each path changes dynamically in data transfer between two nodes, in which plural paths are selectable and when a large round trip delay, not negligible compared with the time constant, exists.

MEANS TO SOLVE THE PROBLEMS

[0022] The present invention is made to solve the

above-mentioned problems. In a first aspect of the present invention, a load distributing method comprises the step of path selection or path selection priority update for a pair of nodes, between which plural communication paths
5 can be selected, upon every packet input to a transmission node, based on path status information on a selectable path, based on identification information on a time from which the path status information is effective or on a transmitted packet, and based on a transmission history
10 after the time from which the path status information is effective or a transmission history after transmission of a packet specified with transmitted packet identification information.

[0023] In a second aspect of the present invention, a
15 load distributing method comprises the steps of monitoring a path status of each path selectable every packet input to a transmission node between two nodes, each which can select plural communication paths, and storing path status information on the path status and a time from which the
20 path status information is effective or packet identification information; estimating a packet arrival prediction time in each path, based on path status information, and a packet transmission history after the time from which the path status information is effective
25 or a packet transmission history after transmission of a

packet specified with the transmitted packet
identification information is transmitted; and the
updating path selection or selection priority, based on
the estimated arrival prediction time.

5 [0024] In the load distributing method of the present
invention, the path status information includes a delay of
a path.

 [0025] In the load distributing method of the present
invention, the path status information includes a
10 transmission rate of a path.

 [0026] In the load distributing method of the present
invention, the path status information includes a load of
a path.

 [0027] The load distributing method of the present
15 invention further comprises the step of correcting a
transmission cost calculation result regarding a packet
transmitted before updating path status information of
each path, when the path status information is updated for
path selection or selection priority update.

20 [0028] The load distributing method of the present
invention further comprises the step of discarding a
history prior to a first packet transmitted on or after a
time from which the latest path status information is
effective, when a transmission cost calculation result of
25 each path is corrected.

[0029] The load distributing method of the present invention further comprises the step of selecting as a packet transmission path a path having an earliest estimation value of a reception completion time at a reception node.

[0030] The load distributing method of the present invention further comprises the step of selecting as a packet transmission path a path having a largest estimation value of a data amount, which can be completely received by a specific time at a reception side.

[0031] The load distributing method of the present invention further interrupts data transmission according to an estimated current path status in each path.

[0032] In the load distributing method of the present invention, a condition for interruption of the data transmission is that an estimated reception completion time is equal to or greater than a specific value.

[0033] In the load distributing method of the present invention, path selection or transmission interruption is determined according to a policy different for each attribute of transmission data.

[0034] The load distributing method of the present invention further comprises the steps of selecting a communications interface corresponding to a destination address of a reception packet from a table which

associates an address of a communications interface with a destination address reachable using the communications interface; selecting a communications interface corresponding to the transmission source address or a communications interface from among the selected communications interfaces when the reception packet has information specifying a transmission source address or a communications interface; selecting a given communications interface from among the selected communications interfaces when the reception packet does not have information on a transmission source address or a communications interface; and sending the reception packet to a selected communications interface.

[0035] Another aspect of the present invention, a node capable of selecting plural packet transmission paths comprises means for updating path selection or selection priority every input packet, based on path status information on each selectable path, based on identification information on time validating the path status information or on a transmitted packet, and based on a transmission history after the time from which the path status information is effective or a transmission history after transmission of a packet specified with transmitted packet identification information.

[0036] A node capable of selecting plural packet

transmission paths comprises monitor means for monitoring
a selectable path status of each path every packet input
at a transmission node and monitoring path status
information on the path status and a time from which the
5 path status information is effective or packet
identification information; memory means for storing the
path status information and a packet transmission history
available after the path status information is validated;
scheduling means for estimating an arrival prediction time
10 of a packet in each path based on the path status
information and based on a packet transmission history
after the path status information is validated and
updating path selection or selection priority based on the
estimated arrival prediction time.

15 [0037] In the node of the present invention, the path
status information includes a delay of a path.

[0038] In the node of the present invention, the path
status information includes a transmission rate of a path.

20 [0039] In the node of the present invention, the path
status information includes a load of a path.

[0040] In the node of the present invention, the
scheduling means corrects a transmission cost calculation
result regarding a packet transmitted prior or updating
when path status information of each path is updated in
25 the updating of path selection or selection priority.

[0041] In the node of the present invention, the scheduling means discards a history before a first transmitted packet validating latest path status information when a transmission cost calculation result of each path is corrected.

[0042] In the node of the present invention, the scheduling means selects as a packet transmission path a path having an earliest estimation value of a reception completion time at a reception node.

[0043] In the node of the present invention, the scheduling means selects as a packet transmission path a path having a largest estimation value of a data amount which can be completely received by a specific time at a reception node.

[0044] In the node of the present invention, the scheduling means interrupts data transmission according to an estimated current path status for each path.

[0045] In the node of the present invention, a condition for interruption of the data transmission is that an estimated reception completion time is equal to or greater than a specific value.

[0046] In the node of the present invention, the scheduling means determines the interruption of path selection or transmission according to a policy different every attribute of a transmission data.

[0047] The node according to the present invention further comprises a table in which an address of a communications interface is associated with a destination address reachable using the communications interface; and
5 routing means for selecting a communications interface corresponding to a destination address of a packet to be transmitted, from the table, selecting a communications interface corresponding to the transmission source address or a communication interface from the selected
10 communications interface when the transmission packet has a information specifying a transmission source address or a communications interface, and sending the transmission packet to a selected communications interface.

[0048] In another aspect of the present invention, a
15 node control program, which is applicable to a node that can select plural packet transmission paths, the node control program controlling the node as means for updating path selection or selection priority every input packet, based on selectable path status information on each path,
20 based on identification information on time validating the path status information or on a transmitted packet, and based on a transmission history after the time from which the path status information is effective or a transmission history after transmission of a packet specified with
25 transmitted packet identification information.

[0049] In another aspect according to the present invention, a node control program, which is applicable to a node that can select plural packet transmission paths, the node control program controlling a node as monitoring means for monitoring a selectable path status of each path for each packet input to a transmission node and monitoring path status information on the path status and identification information on time or packet validating the path status information; and scheduling means for estimating a packet arrival prediction time in each path based on the path status information and based on a transmission history of a packet after the path status information is validated and updating path selection or selection priority based on the estimated arrival prediction time.

[0050] In the node control program of the present invention, the path status information includes a delay of a path.

[0051] In the node control program of the present invention, the path status information includes a transmission rate of a path.

[0052] In the node control program of the present invention, the path status information includes a load of a path.

[0053] The node control program of the present

invention further controls the scheduling means so as to correct a transmission cost calculation result regarding a packet transmitted prior or updating when path status information of each path is updated in the updating of path selection or selection priority.

[0054] The node control program of the present invention further controls the scheduling means so as to discard a history before a first transmitted packet validating latest path status information when a transmission cost calculation result of each path is corrected.

[0055] The node control program of the present invention further controls the scheduling means so as to select as a packet transmission path a path having an earliest estimation value of a reception completion time at a reception node.

[0056] The node control program of the present invention further controls the scheduling means so as to select as a packet transmission path a path having a largest estimation value of a data amount which can be completely received by a specific time at a reception node.

[0057] The node control program of the present invention further controls the scheduling means so as to interrupt data transmission according to an estimated current path status for each path.

[0058] In the node control program of the present invention, a condition for interruption of the data transmission is that an estimated reception completion time is equal to or greater than a specific value.

5 [0059] The node control program of the present invention further controls the scheduling means so as to determine path selection or transmission interruption according to a policy different every attribute of transmission data.

10 [0060] The node control program of the present invention further operates as routing means that selects a communications interface corresponding to a destination address of a packet to be transmitted, from a table in which an address of a communications interface is
15 associated with a destination address reachable using the communications interface, selects a communications interface corresponding to the transmission source address or a communications interface, from the selected
communications interface when the transmission packet has
20 information specifying a transmission source address or a communications interface, and transmits the packet to be transmitted, to the selected communications interface.

 [0061] In another aspect of the present invention, a transmission packet time estimation method for estimating
25 an arrival time or a reception completion time of a

transmission packet, comprises the steps of adding packet
identification information for packet identification to a
transmission packet transmitted from the transmission node
and then transmitting the added information; storing a
5 transmission history of a transmitted packet, together
with packet identification information, at the
transmission node; transmitting information on a path
status from a reception node to a transmitter node,
together with packet identification information on a last
10 packet received in the path status; and estimating an
arrival time or a reception completion time to be
transmitted in the transmission node, based on the path
status and a transmission history of a packet transmitted
subsequent to a packet specified with packet
15 identification information contained in the path status
information.

[0062] In another aspect of the present invention, a
node comprises means for adding packet identification
information for packet identification to a transmission
20 packet and transmitting the added information; means for
recording a transmission history of a transmitted packet,
together with packet identification information; and means
for receiving path status information transmitted from a
reception side and estimating an arrival time or reception
25 completion time of a packet to be transmitted, based on

the path status and based on a transmission history of a packet transmitted subsequent to a packet validating the path status.

[0063] In further another aspect of the present invention, a node control program for a node transmitting packets, the control program operates the node as means for adding packet identification information for packet identification to a transmission packet and then transmitting the added information; means for recording a transmission history of a transmitted packet, together with packet identification information; and means for receiving path status information transmitted from a reception side and estimating an arrival time or reception completion time of a packet to be transmitted, based on the path status and based on a transmission history of a packet transmitted subsequent to a packet validating the path status.

[0064] In the present invention, the transmission history of transmission data (packet) is stored at the time of transmission of data (packet). The transmission history describes, for example, identifiers for identifying transmission data (packets) and transmission start time. The reception side transmits status information (for example, rate or packet delay) of a path through which data (packets) run, to the transmission side.

The path status information contains information for identifying latest data (packets), to which path status information, such as the identifier of the data (packet) described above, is applied above.

5 [0065] The transmission side receives path status information. Using information (for example, identifier) identifying the data (packet) contained in the path status, the transmission side acquires transmitted data (packet) validating the path status referring to the transmission
10 history. As to data subsequent to the data (packet) obtained, the reception completion time is predicted by applying the path status received. Moreover, an arrival time or reception completion time of data to be transmitted in each path is predicted based on the
15 prediction and based on the received status information.

 [0066] In the path selection, an optimum path, for example, a path having a predicted earliest (fastest) arrival time, is selected based on the predicted data arrival time or a data reception completion time.

20

EFFECT OF THE INVENTION

 [0067] According to the present invention, the path status information obtained from a reception side is applied to data (packets) transmitted in past, to which
25 the status information is validated. By reflecting the

result to predict the data (packets) to be transmitted,
the arrival time or reception completion time of the data
is predicted. This method can predict with higher accuracy,
compared with the conventional art. The reason is that, in
5 the conventional art, the path information obtained from
the reception side is applied only to data (packets) to be
transmitted to predict the arrival time or reception
completion time.

[0068] According to the present invention, in order
10 to transfer data between two nodes, each which can select
plural paths, the path status information obtained from
the reception side is applied to data transmitted in past
which validates the information. The result is reflected
to predict data to be transmitted and the data arrival
15 time or the reception completion time is predicted. Since
the path selection is performed based on the prediction, a
suitable path is selected, without selecting an erroneous
path due to a prediction error. Thus the problem that the
multiplexing efficiency decreases can be avoided.

[0069] According to the present invention, in order
20 to transfer between two nodes, which can select plural
paths, the path resources can be effectively used even
when the performance of each path dynamically changes and
a large return delay, not negligible, exists, compared
25 with the time constant.

BRIEF DESCRIPTION OF THE DRAWINGS

[0070]

5 [Fig. 1] Fig. 1 is a diagram illustrating a path
between nodes, according to the present invention.

[Fig. 2] Fig. 2 is a diagram illustrating the
configuration of a transmitter node.

10 [Fig. 3] Fig. 3 is a diagram explaining the timing of
a packet process at a transmitter node and the timing of a
packet process at a receiver node, according to an
embodiment of the present invention.

[Fig. 4] Fig. 4 is an operational flowchart for a
scheduler according to an embodiment.

15 [Fig. 5] Fig. 5 is a diagram explaining the timing of
a packet process at a transmitter node and the timing of a
packet process at a receiver node, according to another
embodiment of the present invention.

[Fig. 6] Fig. 6 is an operational flowchart for a
scheduler according to another embodiment.

20 [Fig. 7] Fig. 7 is a diagram explaining prediction of
an arrival time or a reception completion time at a
transmitter node and prediction of an arrival time or a
reception completion time at a receiver node, according to
an embodiment.

25 [Fig. 8] Fig. 8 is a diagram explaining a

conventional art.

[Fig. 9] Fig. 9 is a configuration example where a communications interface has an individual IP address.

5 [Fig. 10] Fig. 10 is a diagram illustrating the internal configuration of the transmission node 100-2 in the system of Fig. 9.

[Fig. 11] Fig. 11 is a diagram illustrating the IP rooting processor 1318.

10 [Fig. 12] Fig. 12 is a diagram illustrating the path control table 1321.

[Fig. 13] Fig. 13 is a diagram illustrating the IF address management table 1322.

15 [Fig. 14] Fig. 14 illustrates the path to communications interface correspondence table 1330, to be saved in the memory 315-2.

[Fig. 15] Fig. 15 shows the configuration of the transmission IP packet creation data 1510.

[Fig. 16] Fig. 16 shows the configuration of the delivery IP packet 1610.

20 [Fig. 17] Fig. 17 is a flowchart for transmission IF decision of the transmission IF decision/delivery processor 1320.

[Fig. 18] Fig. 18 is a diagram illustrating other transmission IP packet creation data.

25 [Fig. 19] Fig. 19 is a diagram illustrating other

delivery IP packet.

[Fig. 20] Fig. 20 is a flowchart for other transmission IF decision of the transmission IF decision/delivery processor 1320.

5 [Fig. 21] Fig. 21 is a diagram illustrating entries of the path control table 1321.

EXPLANATION OF SYMBOLS

[0071]

| | |
|----|------------------------------|
| 10 | 10 Data creation node |
| | 11 Destination node |
| | 100 Transmission node |
| | 101 Reception node |
| | 102 Cable network |
| 15 | 200 Radio transmitter |
| | 201 Radio receiver |
| | 202 Radio link |
| | 300 Radio network |
| | 310 Communications interface |
| 20 | 312 Queuing section |
| | 313 Scheduler |
| | 314 Path status monitor |
| | 315 Memory |

25 **BEST MODE FOR CARRYING OUT THE INVENTION**

[0072] A best mode for carrying out the present invention will be explained below.

[0073] It is assumed that an aspect of the present invention is applied to the technique combining plural
5 paths, of which the rate as well as the delay vary dynamically, particularly, paths including radio links to construct one logical path. Inverse multiplexing will be explained below as an example of a technique of constructing one logical path through a combination of
10 plural paths.

[0074] Fig. 1 shows a structural example of the present embodiment.

[0075] Referring to Fig. 1, a transmission node 100 and a reception node 101, each implementing inverse
15 multiplexing, exist in a path between a data creation node 10 and a destination node 11. There are three paths between the transmission node 100 and the reception node 101. Radio links 202-1 to 202-3 are respectively in communication paths between radio transmission means 200-1
20 to 200-3 and radio reception means 201-1 to 201-3. Three paths are shown in Fig. 1 but two paths or more may be disposed. All paths between the transmission node and the reception node may be constructed by radio. However, the paths generally include cable networks 102. Generally each
25 radio link belongs to a different radio network 300. In

this example, the path 202-1, 202-2 is a Cellular network 300-1 and the radio network 300-2 is a radio LAN.

[0076] In the system of Fig. 1, the transmission node 100 distributes traffics received by the data creation
5 node 10 to respective paths, based on the status information. The reception node 101 reintegrates the traffics sent from the transmission node 100 via respective paths and transmits the result to the destination node 11.

10 [0077] The internal configuration of the transmission node 100 is shown in Fig. 2.

[0078] The communication interface 310-1 receives traffics transmitted from the data creation node 10 toward the destination node 11. The transmission side
15 communication interface 310-2 or 310-3 in an inverse multiplexing link transmits the traffics via the queuing section 312 and the scheduler 313. In some cases, plural paths share the physical link closer to the transmission node. Hence, the paths constructing the inverse
20 multiplexing link do not necessarily correspond to the communications interfaces one by one.

[0079] The scheduler 313 takes input data out of the queuing section 312 and sends it to a specific path. The path selection for the transfer of the captured data is
25 implemented by referring to the status of the path managed

by the path status monitor 314. The path status monitor 314 intermittently receives status information of each path and information identifying a transmission packet validating updating (hereinafter referred to as a report),
5 from the reception node 101 via the communications interface 310-2 or 310-3. Thus, the path status monitor 314 updates path status information stored in the memory 315 based on them.

[0080] The path status information generally
10 represents information on a communication performance indicator of communication performance. The present embodiment uses a path rate and a packet delay in path status information. An effective time of path status information does not means updating of path status
15 information in the transmission node but means the time when a path becomes a path status shown with the path status information or a time specified with a packet transmitted in the path status. Hence, a transmission history after the effective time of path status
20 information means a packet transmission history transmitted after the effective time of path status information. The time is obtained based on packet information (identifier identifying a packet) originally used for the time measurement or based on a reception time
25 transmitted from the reception node. However, the time may

be varied somewhat. Various methods in which the reception node measures the rate and delay to obtain path status information have been proposed. The method assumed in the present embodiment will be described below.

5 [0081] The transmission node 100 inserts an identifier and a transmission time in each packet transmitted to the reception node 101 and then transfers the combined data. The reception node measures the packet delay by comparing the transmission time inserted in the
10 transmission node with the time when the reception node itself has received the packet. The transmission node periodically transmits a train of packets for measurement and the reception node can estimate the rate based on variations of the arrival times. The estimation method is
15 disclosed in detail, in for example, "What DoPacket Dispersion Techniques Measure?" written by Dovrolis, Ramanathan, and Moore, IEEE INFOCOM 2001. In the document by Dovrolis et al., the transmission node simultaneously
20 transmits two packets and the transmission side estimates the link rate based on a difference between arrival times of two packets. The transmission delay spreads a packet arrival interval. Because the transmission delay relates to a link rate, the rate can be estimated by the arrival interval.

25 [0082] The reception node 101 periodically transmits

those measurement values as path status information to the transmission node 100. At the same time, the reception node 101 transmits the identifier of the latest packet identifier received by that time as packet identification information validating status information to be
5 transmitted. The transmission node receives those sets of information as reports. This system has been shown as an example. The enabling possibility of the invention does not depend on the method of determining and transmitting
10 path status information and a packet validating the path status information.

[0083] As to the packet to be next transmitted, the scheduler 313 refers to a transmission history available after current path information and a packet validating the
15 information are transmitted for each transmission path. Thus, the scheduler 313 predicts a delay of the packet arriving at the reception node 101. The transmission history is stored in the memory 315. The scheduler 313 selects a path with a predicted minimum arrival delay as a
20 transmission path for a packet to be next transferred and adds the transfer time to the transmission history in the memory 315 after the packet transfer to the selected path.

[0084] Fig. 3 illustrates an arrival delay estimation method for each path operated with the scheduler 313. In
25 Fig. 3, numeral 400-1, 400-2, or 400-3 represents a data

packet. The transmission history and prediction in the transmission node 100 and the reception history and prediction in the reception node 101 are shown on the time axes, respectively. The transmission node, for example, starts transmitting the data packet 400-1 at time T1 and then ends its transmission operation at time T2. Moreover, the reception node starts receiving the packet 400-1 at time T3 and then completes its reception operation at time T4. The difference I1 between time T1 and time T3 corresponds to a transmission delay. The difference I2 between time T4 and T2 corresponds to a total delay, which is the transmission delay I1 plus a packet dispersion occurring due to the rate difference between the transmission interface and the transfer path.

[0085] Here, it is assumed that the packet 400-3 is transmitted at the current time being the point TP on the time axis. As to the path, the path status information is updated because a report result report is received between T5 and TP. The path status information is validated from the packet 400-1. The reception completion time of the packet 400-3, in consideration of the packets 400-1 and 400-2, is estimated based on the transmission history of the packet (400-1, 400-2) validating the path status information.

[0086] The reception start time and the completion

time of the packet 400-2 transmitted behind the packet 400-1, in the reception node 101, are estimated based on the rate and the transmission delay containing path status information currently being received. In Fig. 3, the
5 estimated reception start time is T5 and the reception completion time is T7. If the delay time shown with the path status information is equal to I1, the packet 400-3 transmitted at TP is to be received at time T6. However, it is assumed that the reception node has not yet
10 completely received the packet 400-2. Hence, the estimated reception start time of the packet 400-3 is assumed to be T7 at which the packet 400-2 is completely received. The estimated reception completion time of the packet 400-3 becomes T8, to which the packet dispersion estimated from
15 the path rate, contained in the path status information, is added. Similarly, the reception completion time of the packet 400-3 is estimated for each path. The packet 400-3 is transmitted to a path providing a fastest reception time.

20 [0087] The path status information used for arrival time estimation of the reception side at time TP in Fig. 3 is updated upon notification of a report result between T5 and TP. The arrival time of the packet 400-2 transmitted is estimated based on old path status information at TP.
25 It is assumed that the old status information is

information A and that new information updated between T5 and TP is information B. If the delay and path rate contained in the information A and B differ from each other due to variations of the link status, the arrival time estimation based on the information A will differ from the arrival time estimation based on information B shown in Fig. 3. Therefore, the arrival time prediction of the packet 400-1 and the arrival time prediction of the packet 400-2 shown in Fig. 3 reflect on correction of the result obtained with the information B. When the path status information is once updated, the transmission histories of packets prior to the packet validating the updating are discarded because they are unnecessary for reference.

[0088] Moreover, a difference between the packet arrival time estimation and reception completion time in the present invention and the packet arrival time estimation and reception completion time in the conventional art will be specifically explained below by referring to Fig. 7.

[0089] Referring to Fig. 7, numeral 500-1, 500-2, or 500-3 represents a data packet to be transmitted. The transmission history (in solid line) of a packet in the transmission node and the reception history (in solid line) of an actual packet in the reception node are

indicated on the time axes, respectively. The prediction (in dotted line) of a packet arrival time and a reception completion time are indicated on the time axes. On the upper portion of Fig. 7, the curve shows a rate change of a transmission path. Fig. 7 shows the status where the transmission rate delays with an elapse of time.

[0090] In actual transmission and reception of packets, the transmission node first starts transmitting the data packet 500-1 at time T1 and terminates the transmission at time T2. The reception node starts receiving the data packet 500-1 at time T3 and terminates the reception at time T5.

[0091] Thereafter, the transmission node starts transmitting the data packet 500-2 at time T4 and terminates its transmission at time T8. The reception node starts receiving the data packet 500-2 at time T7 due to a decrease in communication rate and terminates the reception at time T13.

[0092] Moreover, the transmission node starts transmitting the data packet 500-3 at time T9 and terminates the transmission at time T10. The reception node starts receiving the data packet 500-3 at time T14 and terminates the reception at time T16.

[0093] Next, prediction of a packet arrival time and a reception completion time according to the present

invention in the actual packet transmission and reception will be explained below.

5 [0094] The transmission node starts transmitting the data packet 500-1 at time T1, together with information on the transmission time and the packet identifier of the data packet 500-1, and terminates its transmission at time T2. The transmission node stores as a transmission history the transmission time and the packet identifier of the data packet 500-1.

10 [0095] In succession, the transmission node starts transmitting the data packet 500-2 at time T4, together with the information on the transmission time and the packet identifier of the data packet 500-2 and terminates the transmission at time T8. At this time, in a manner
15 similar the above-mentioned method, the transmission node stores as a transmission history the transmission time and the packet identifier of the data packet 500-2.

20 [0096] The reception node starts receiving the data packet 500-1 at time T3 and terminates the reception at time T5. At this time, the reception node transmits as a report information on the identifier and the reception completion time and the communication rate of the data packet 500-1, to the transmission node.

25 [0097] If receiving a report from the reception node at time TX, the transmission node decides the transmission

time T1 of the data packet 500-1 at the time the path status information of the report validates, based on the identifier of the data packet 500-1 contained in the report.

5 [0098] The transmission node predicts the arrival time and the reception completion time of the data packet 500-3, based on the transmission history after the time T1. In the transmission history after the time T1, the data packet 500-1 and 500-2 are to be handled. According to the
10 report, the reception completion time of the data packet 500-1 is time T5. The arrival time and the reception completion time of the data packet 500-2 are used to predict the communication rate and the reception completion time of the data packet 500-1 in the report.
15 Based on those sets of information, the transmission delay or the total packet delay time is known. The transmission delay is the difference I1 between time T1 and time T3. The total delay of a packet is the transmission delay I1 plus the packet dispersion caused by the rate difference
20 between the transmission interface and the transfer path. The difference between time T5 and time T2 is I2. The arrival time of the data packet 500-2 is obtained based on the transmission delay I1 and is predicted to be the time T6, which is the transmission time T4 of the data packet
25 500-2 plus the transmission delay I1. Since the packet

delay can be obtained based on the communication rate in the report, the reception completion time of the data packet 500-2 is predicted to be time T12. Hence, in this path, the arrival time of the data packet 500-3 in the reception node can be predicted to be a time after the time T12. When the reception starts from the time T12, the reception completion time can be predicted to be time T15 through the packet dispersion prediction. According to the prediction results, the difference between the arrival time T12 of the predicted data packet 500-3 and an actual arrival time T14 of the data packet 500-3 is time ID1.

[0099] In the prediction of an arrival time and a reception completion time of a packet in the conventional art, the path status shown in a report obtained by the reception node is validated at the report arrival time of the report. The path status is applied, without any change, to predict the data packet 500-3. That is, the arrival time of the data packet 500-3 is predicted to be the time T11, which is the transmission time T9 of the data packet 500-3 plus the transmission delay I1. In such a prediction result, the difference between the arrival time T11 of the expected data packet 500-3 and the arrival time T14 of the actual data packet 500-3 is time ID2.

[0100] As apparent from the drawing, as to the error time ID1 and the error time ID2 with respect to an actual

arrival time, the error time ID1 predicted according to the present invention is smaller than the error time ID2.

[0101] Next, the procedure of selecting the path of a scheduler, containing the above-mentioned arrival time estimation, is shown in Fig. 4.

[0102] First, the queuing section 312 receives packets (step 100).

[0103] Next, after the last packet was transmitted through any one of paths, it is decided whether or not the path status information has been updated (step 101). When the path status information is updated, the transmission histories existing before the path status information is validated are erased in all the updated paths (step 102).

[0104] In succession, in each path, the packet arrival time is predicted (estimated) based on the path status information and the transmission history (step 103). Thus, a packet is transmitted to the path a predicted (estimated) fastest arrival time (step 104).

[0105] Finally, the transmission history of the path used for packet transmission is updated (step 105).

[0106] As described above, in the updating of path status information, the predicted arrival time of a packet transmitted before the updating is corrected. The corrected arrival time is reflected on judgement of the subsequent packet transmission. Accordingly, the past

transmission record can be compensated. The effect of the compensation becomes significant when the round delay of each path result is larger and cannot be ignored to the cycle of the path status variation. The reason will be described below.

[0107] Where the path status variation occurs in the time such as a round delay, the status of the corresponding path may have already changed when the transmission node captures a set of status information. Hence, such information is unreliable. Therefore, it is impossible to optimize the path selection and the timing setting at the time of packet transmission. Generally, the packet is transmitted with an unsuitable path and with an unsuitable timing. Correcting the arrival time prediction upon the status information updating is equivalent to estimating the already performed unsuitable transmission impact from a lapse of the interval between status information updating operations. For example, when the packet has been transmitted at an excessive rate, the arrival prediction time of a transmitted packet is extended through the updating of the status information. This increases the path transmission cost.

[0108] As described above, the correction of the arrival time prediction of a past transmission packet due to the updating of status information effectively

optimizes the path selection. However, when the correction of the arrival time prediction is fed back for the transmission timing control, in addition to the path selection, the congestion control of each path is
5 optimized in the long term.

[0109] Next, another embodiment will be explained below.

[0110] An operation of the following embodiment embodying the simple timing control will be explained
10 below.

[0111] In a manner similar to that of the above-mentioned embodiment, another embodiment estimates a reception completion time of a transmission packet for each path and selects a path having a highest evaluation
15 value. However, the embodiment can introduce the simple timing control, under which an allowable estimation delay is defined for a new path and the transmission node 100 controls the transmission timing such that the estimation delay does not exceed the allowable estimation value. The
20 operation of the embodiment will be explained by referring to Fig. 5.

[0112] In Fig. 5, the allowable estimation delay TM means that it must be estimated that a packet is completely received by TM+TP to transmit the packet at TP.
25 The reception completion time of the packet 400-3

estimated with the means, identical to that in the above embodiment, becomes $T8$ corresponding to a future value with respect to $(TM+TP)$. The transmission node 100 cannot transmit the packet 400-3 out of the path until the
5 estimated reception completion time becomes $TM+TP$. In this case, the transmission node 100 retains the packet 400-3 until the estimated reception completion time of any one of paths becomes $TM+TP$ or less. The transmission node 100 transmits the packet through the path of which the
10 retention is released at an earliest time. The operation flow of the scheduler 313 in that mode is shown in Fig. 6.

[0113] First, the queuing section 312 receives a packet (step 200).

[0114] Next, it is decided whether or not the path
15 status information updated after the last packet has been transmitted via any one of paths (step 201). When the path status information is updated, transmission histories available before the path status information is validated are erased in all the updated paths (step 202).

[0115] In succession, the packet arrival time is
20 predicted (estimated) based on the path status information and the transmission history in each path (step 203). Using the predicted (estimated) arrival time, it is decided whether or not a predicted (estimated) arrival
25 time $>$ (a current time + an allowable estimation delay) in

all paths (step 204). In the case of (predicted
(estimated) arrival time) \leq (current time + allowable
estimation delay) in any one of paths, the packet is
transmitted to the path with an earliest predicted
5 (estimated) arrival time of the paths (step 205).

[0116] When (predicted (estimated) arrival time) $>$
(current time + allowable estimation delay) is held in all
paths, the status is waited until the current time becomes
(predicted arrival time - allowable estimation delay) in
10 any one paths. When the conditions are satisfied in any
one of paths, the flow proceeds to the step 205 (step 206).

[0117] Finally, the transmission history of the path
used for packet transmission is updated (step 207).

[0118] The allowable delay value TP may be set
15 independently for each path. For example, when the delay
of each path or the buffer amount of a server through
which the packet pass is largely different, a different TM
set value may be set in each path, particularly, under a
heavy load. Thus, it is considered that the band of each
20 path can be effectively used.

[0119] If a packet missing ratio and a line use
charge, for example, can be monitored in addition to an
estimated reception completion time, they may be estimated
preferentially. The decision method may depend on the data
25 attribute. For example, the selection is performed which

places greater emphasis on delay for voice data or on the line use charge for file transfer data not urgent. In any case, the present invention is characterized by capturing a transmission packet or time validating the updating at the same time when the transmission node updates path status information, estimating an impact of the transmission cost with the transmission history available after the effective packet or effective time, and transmitting the next packet to a cost minimized path. As a result, even when the delay is large, not negligible to the time constant of path status variation, the impact to cost standard already provided through the past unsuitable transmission can be reflected on the adjustment the subsequent transmission timing to compensate the delay. Thus, the path use efficiency is effectively improved.

[0120] Next, the path selecting method of the present invention, applied to the IP network as shown in Fig. 9, will be explained below.

[0121] Fig. 9 shows the configuration of the communications interface having an individual IP address, in the present embodiment.

[0122] Referring to Fig. 9, the transmission node 100-2 and the reception node 101-2, each having plural paths, are between the data creation node 10-2 and the address node 11-2. Three paths are between the

transmission node 100-2 and the reception node 101-2. In the configuration of Fig. 9, the transmission node 100-2 includes communications interfaces S-IF#1 1310-1, S-IF#2 1310-2, S-IF#3 1310-3, and S-IF#4 1310-4, acting as
5 transmission means. The communications interfaces S-IF#1 1310-1 to S-IF#3 1310-3 belong to radio networks 1300-1 to 1300-3, respectively. The communications interfaces S-IF#1 1310-1 to S-IF#3 1310-3 are connected to the carrier network gateways 1400-1, 1400-2 and 1400-3, which are
10 installed in the cellar carrier network, via the radio links 202-4 to 202-6, respectively. The communications interface S-IF#4 1310-4 is connected to the communications interface D-IF 1200, installed in the data creation node 10-2, via the cable link or radio link 202-7. The
15 reception node 101-2 includes as reception means the communications interface R-IF 1500-1. The reception node 101-2 is connected to the carrier network gateways 1400-1 to 1400-3 via the communications interface R-IF 1500-1 and the cable network 102-2.

20 [0123] In the system of Fig. 9, the communications interfaces S-IF#1 1310-1 to S-IF#3 1310-3 use Point-to-Point Protocols (PPP) to allocate [100.1.2.3], [110.1.2.3] and [120.1.2.3] as individual IP addresses, respectively. Moreover, the communications interface R-IF 1500-1
25 includes fixedly setting means or DHCP means to allocate

[200.7.8.9] as an IP address. The communications interface D-IF 1200 includes fixedly setting means or DHCP means to allocate [192.168.2.50] as an IP address.

[0124] Fig. 10 illustrates the internal configuration of the transmission node 100-2 in the system of Fig. 9. For the IP communications, the transmission node 100-2 includes a packet queuing section 312-2, a scheduler 313-2, a path status monitor 314-2, a memory 315-2, an IP packet creator 1316, an IP packet queuing section 1317, and an IP routing processor 1318.

[0125] The scheduler 313-2 derives input data (destination data and transmission data) from the queuing section 312-2 to select a specific communications interface. The communications interface (path) used to transfer the derived data is selected by referring to the path status managed by the path status monitor 314-2. In the path selection method by the scheduler 313-2 and the path status monitor 314-2, the path is selected and decided, in a manner similar to that by the scheduler 313 and the path status monitor 314. The memory 315-2 stores the path status information, in a manner similar to that of the memory 315. The following explanation is made as to the scheduler 313-2 and the path status monitor 314-2, which select the communications interface (path), in the operation similar to the above-described operation.

[0126] The IP packet creator 1316 adds an IP header to data received from the scheduler 313-2 to create an IP packet. The IP packet queuing section 1317 is a buffer for storing an IP packet created by the IP packet creator 1316. The IP routing processor 1318 derives an IP packet stored in the IP packet queuing section 1317, decides the next transfer destination of the IP packet, and transmits it using a suitable one of the communications interfaces S-IF#1 1310-1 to S-IF#4 1310-4.

[0127] Fig. 11 is a diagram representing the configuration of the IP routing processor 1318. The IP routing processor 1318 has a path control table 1312 (shown in Fig. 12) used to decide the transfer destination of an IP packet handed over. Moreover, the IF address management table 1322 (shown in Fig. 13), which records IP addresses allocated to installed communications interfaces, is prepared. The path control table includes a transmission IF decision/transmission processor 1320. Using the path control table 1321 and the IF address management table 1322, the transmission IF decision/transmission processor 1320 decides the transfer destination of an IP packet handed over to the IP routing processor 1318 and then transmits the IP packet to the corresponding communications interfaces S-IF#1 1310-1 to S-IF#4 1310-4.

[0128] Fig. 14 depicts an example of a table storing correspondences between communication paths and the corresponding communications interfaces, which are stored in the memory 315-2, (hereinafter referred to as a path to communications interface correspondence table 1330).

5 Communication path numbers, or identifiers, for respective communication paths, recognized by the scheduler 313-2 and the path status monitor 314-2, and communications interface names used for communication path connection are

10 stored in the entries of the path to communications interface correspondence table 1330.

[0129] In the path control table 1321 shown in Fig. 12, a network mask and a gateway IP address representing the next transfer destination to send an IP packet to the

15 network are generally described for each destination network IP address being an IP packet transmission destination. The path control table 1321 also records information on a communications interface connected to the sub-network having the same address as that of the gateway

20 IP address. Actually, a multicast address may be often set to transmit packets to all communication equipment belonging to the same sub-network. However, the illustration and explanation will be omitted here.

[0130] The IF address management table 1322, shown in

25 Fig. 13, records correspondences between communications

interfaces installed in the transmission node and IP addresses allocated to them. The content of the IF address management table 1322 is updated every time the IP address allocated to each communications interface is changed.

5 [0131] In the present embodiment, the reception node 101-2 includes as reception means the communications interface R-IF 1500-1. The transmission node 100-2 includes the communications interfaces S-IF#1 1310-1 to S-IF#4 1310-4, each to which a different IP address is
10 allocated. For that reason, the path control table 1321 has entries 1321-a, 1321-b and 1321-c, which register different gateway IP addresses and different communications interfaces, respectively, to the same destination network IP address of [200.7.8.9]. Moreover,
15 the path control table 1321 has the path information 1321-d being path information to the data creation node 10-2.

 [0132] The transmission IF decision/transmission processor 1320 implements the so-called IP routing process that decides an interface used to transmit an IP packet,
20 using the header information of the received IP packet and the path control table 1321.

 [0133] The present embodiment, as shown in Fig. 12, has plural entries in which "destination network IP networks" are overlapped. The IP routing process decides
25 an interface used for delivery.

[0134] Thereafter, in the present embodiment shown in Fig. 9, the scheduler 313-2, for example, selects a transmission path having the communication path number 3, registered in the path to communications interface correspondence table 1330. Thus, the IP packet is transmitted using the communications interface S-IF#2 202-5 corresponding to the decided path. Selecting other path can be realized in the same process.

[0135] When selecting the communication path of the communication path number 3, the scheduler 313-2 acquires the communications interface S-IF#2 corresponding to the communication path from the path to communications interface correspondence table 1330.

[0136] Next, the scheduler 313-2 retrieves the IF address management table 1322 within the IP routing processor 1318 as the captured communications interface S-IF#2 and then obtains the IP address [110.1.2.3] allocated to the communications interface S-IF#2. Next, the scheduler 313-2 creates the transmission IP packet creation information (hereinafter, referred to as transmission IP packet creation data) and then hands it over the IP packet creator 1316.

[0137] Fig. 15 shows the configuration of the transmission IP packet creation data 1510. The creation data 1510 stores as "destination IP address" the IP

address [200.7.8.9] of the communications interface 1500-1 of the reception node 101-2 and stores as "transmission source IP address" the IP address [110.1.2.3] of the communications interface S-IF#2 1310-2.

5 [0138] The IP packet creator 1316 creates the transmission IP packet 1610, shown in Fig. 16, based on the received transmission IP packet creation data 1510, and adds it to the IP packet queuing section 1317. Based on the information contained in the transmission IP packet
10 creation data 1510, the transmission IP packet 1610 stores the IP address [110.1.2.3] of the communications interface S-IF#2 1310-2 in the "transmission source IP address" field of the IP header. Moreover, the transmission IP packet 1610 stores the IP address [200.7.8.9] of the
15 communications interface R-IF 1500-1 of the reception node 101-2 in the "destination IP address" field of the IP header.

 [0139] Next, in the IP routing processor 1318, the transmission IF decision/delivery processor 1320 derives
20 the transmission IP packet 1610 from the IP packet queuing section 1317 and decides the communications interface handing over the transmission IP packet, using the path control table 1312 and the IF address management table 1322. Thus, the transmission process is executed.

25 [0140] Fig. 17 is a flowchart showing the flow of

transmission IF decision of the transmission IF
decision/delivery processor 1320 in the present embodiment.
When the path control table 1321 is constructed with
entries of x, the transmission IF decision/delivery
5 processor 1320 prepares as variables a count value i, a
first candidate list, a count value m, a count value j,
and a second candidate list, and a count value n. The
count value i represents an entry in the path control
table 1321 processed when a retrieval process is performed
10 under the first condition. The first candidate list stores
an arrangement of entries in the path control table
selected as a transmission candidate under the first
condition. The count value m represents the number of
entries stored in the first candidate list. The count
15 value j represents an entry in the first candidate list
processed when a selection process is performed under the
second condition. The second candidate list stores an
arrangement of transmissions interfaces selected as a
transmission candidate under the second condition. The
20 counter value n represents the number of entries stored to
the second candidate list. First, the transmission IP
packet 1610 is derived from the IP packet queuing section
1317 (step 300). "1" is set to the count values i and j as
an initial value and "0" is set to the count values m and
25 n (step 301). Then the flow goes to the step 302.

[0141] Next, by comparing the count value i with the count value x , it is decided whether or not the first candidate retrieval has been completed (step 302).

Specifically, when the count value i is equal to the count value x , the first candidate retrieval has completed. When the count value i is not equal to the count value x (NO in step 302), the value stored in the destination IP address of the transmission IP packet 1610 is compared with the value stored in the destination network IP address of the i -th entry of the path control table (step 303). When the comparison result indicates a match (YES in step 303), the content of the i -th entry of the path control table is copied to the m -th entry of the first candidate list (step 304). Thus, the count value m increments by one (step 305) and the count value i increments by one (step 306). The flow goes to the next path control table entry process. When the value stored in the destination IP address of the transmission IP packet 1610 does not match the value stored in the destination network IP address of the i -th entry of the path control table (NO in step 303), the flow goes to the step 306 without any change.

[0142] By repeating the process between the step 302 and the step 306, the process is performed to respective entries sequentially from the first entry of the path control table. When the process reaches the last entry

(YES in step 302), the flow goes to the step 307 to decide whether or not entries exist in the first candidate list.

[0143] Specifically, the presence or absence of the first candidate is determined by the count value m. When the count value m is "0" (YES in step 307), the communications interface for transmitting the transmission IP packet 1610 does not exist, so that the process ends. When the count value m is not "0" (NO step 307), the first candidate exists, so that the process goes to the step 308.

Next, whether or not the count value m is "1" is determined (step 308). When the count value m is "1" (YES in step 308), the communications interface for transmitting the transmission IP packet 1610 is uniquely determined. Hence, the communications interface recorded to the first entry in the first candidate list is captured (step 309). The transmission IP packet 1610 is transmitted using the captured communications interface (step 322) and thus the process ends. When the count value m is not "1" (NO in step 308), the communications interface used to transmit the transmission IP packet 1610 cannot be determined uniquely, so that the flow goes to the step 310.

[0144] In the step 310, it is decided whether or not the second candidate retrieval has completed. Specifically, when the count value j is equal to the count value m, the second candidate retrieval has completed. When the count

value j is not equal to the count value m (NO in step),
the communications interface of the j-th entry in the
first candidate list is first captured (step 311). Next,
the allocation IP address corresponding to the
5 communications interface captured in the step 310 is
acquired from the IF address management table 1322 (step
312). In succession, the transmission source IP address of
the transmission IP packet 1610 is compared with the
allocation IP address captured in the step 311 (step 313).
10 When the transmission source IP address of the
transmission IP packet 1610 is matched with the allocation
IP address (YES in step 313), the communications interface
captured in the step 311 is copied to the n-th entry in
the second candidate list (step 314). Thus, the count
15 value n increments by one (step 315). Moreover, the count
value j increases by one (step 316). The flow goes to the
process of the next entry in the first candidate list.
When the transmission source IP address of the
transmission IP packet 1610 is not matched with the
20 allocation IP address (NO in step 313), the flow goes to
the step 316 without any change.

[0145] By repeating the process between the steps 310
and 316, each entry process is performed sequentially from
the first entry in the first candidate list. When the
25 process reaches the final entry (YES in step 310), the

flow goes to the step 317 to determine whether or not entries exist in the second candidate.

[0146] Specifically, the existence or absence of the second candidate is determined by the count value n. The count value (n =) "0" means that the communications interface capable of transmitting the transmission IP packet 1610 to the first candidate list exists but is not decided uniquely. The count value (n =) "1" means that the communications interface capable of transmitting the transmission IP packet 1610 is uniquely decided. The count value (n =) "2" or more means that the communications interface capable of transmitting the transmission IP packet 1610 to the second candidate list exists but is not uniquely decided.

[0147] When the count value n is "1" (YES in step 317), the communications interface is captured from the first entry in the second candidate list (step 318). Using the captured communications interface (step 322), the transmission IP packet 1610 is transmitted and the process ends. When the count value n is not "1" (NO in step 317), the flow goes to the step 319.

[0148] Next, when the count value n is "0" (YES in step 319), the transmission interface is captured from a given entry in the first candidate list (step 320). Using the captured transmission interface, the transmission IP

packet 1610 is transmitted (step 322) and the process ends.
When the count value n is not "0" (NO in step 319), the
transmission interface is captured from a given entry in
the second candidate list (step 321). Using the captured
transmission interface, the transmission IP packet 1610 is
transmitted (step 322) and the process ends.

[0149] In the IP header of the transmission IP packet
1610, [200.7.8.9] is stored to the destination IP address
and [110.1.2.3] is stored in the transmission source IP
address. Therefore, the first candidates correspond to the
entries 1321-a, 1321-b and 1321-c in the path control
table 1321, shown in Fig. 21. In the execution of the
process in the step 307, the first entry of the first
candidate list stores the content of the entry 1321-a of
the path control table 1321. Moreover, the second entry of
the first candidate list stores the content of the entry
1321-b of the path control table 1321. The third entry of
the first candidate list stores the content of the entry
1321-c of the path control table 1321. "3" is stored for
the count value m. In the step 312, for j=1, [100.1.2.3]
is captured as an allocation IP address and for j=2,
[110.1.2.3] is captured as an allocation IP address and
for j=3, [120.1.2.3] is captured as an allocation IP
address. In the execution of the step 317, the second
candidate list stores "S-IF#2 only as a transmission

interface. "1" is stored as the count value. That is,
based on the decision result by the process of the step
317, the process of the step 318 is executed. In the step
318, "S-IF#2 is captured as a transmission interface. The
5 transmission IF decision/delivery processor 1320 transmits
the transmission IP packet 1610, using the communications
interface S-IF#2 1310-2 corresponding to the path of the
path number 3 selected by the scheduler 313-2.

[0150] As a result, the transmission IP packet 1610
10 is transmitted from the communications interface S-IF#2
1310-2. The communications interface R-IF 1500-1 in the
reception node 101-2 receives the transmission IP packet
1610 via the carrier network GW 1400-3.

[0151] The present embodiment includes plural
15 communications interfaces, in which the transmission node
100-2 and the reception node 101-2 are not in a one-to-one
correspondence, each communications interface being
allocated with an individual IP address. By implementing
the above-mentioned procedure, even such a configuration
20 can transmit the IP packet using the path decided by the
scheduler 313-2. The IP packet, which does not
particularly store a transmission source IP address and
wants the IP routing in the conventional art, can be
subjected to the IP routing process.

25 [0152] Moreover, according to another embodiment, the

scheduler 313-2 in the transmission node 100-2 may select a communication path of the communication path number "3" and then converts the transmission IP packet creation data handed over to the IP packet creator 1316 into the format shown in Fig. 18.

[0153] Fig. 18 shows the transmission IP packet creation data 1511 in the present embodiment. The scheduler 313-2 stores, as "destination IP address" of the transmission IP packet creation data 1511, the IP address [200.7.8.9] of the communications interface 1500-1 in the reception node 101-2. Moreover, the scheduler 313-2 stores, as "transmission source interface" of the transmission IP packet creation data 1510, the communications interface "S-IF#2" corresponding to the communication path number "3" captured from the path to communications interface correspondence table 1330.

[0154] When receiving the transmission IP packet creation data 1511, the IP packet creator 1316 creates the transmission IP packet 1611 shown in Fig. 19 and adds it to the IP packet queuing section 1317. In the transmission IP packet 1611, the communications interface "S-IF#2" is stored in the "transmission source IP address" field of the IP header, based on the information contained in the transmission IP packet creation data 1511. Moreover, the IP address [200.7.8.9] of the communications interface R-

IF 1500-1 in the reception node 101-2 is stored in the "destination IP address" field of the IP header.

[0155] In the IP routing processor 1318, the transmission IF decision/delivery processor 1320 captures the transmission IP packet 1611 from the IP packet queuing section 1317 and decides the communications interface handing over the transmission IP packet, using the path control table 1321 and the IF address management table 1322. Thus, the transmission process is executed.

[0156] Fig. 20 is flowchart showing the transmission IF decision of the transmission IF decision/deliver processor 1320 in the present embodiment. In this embodiment, when the path control table 1321 is constructed with entries of x, the transmission IF decision/delivery processor 1320 prepares as variables a count value i, a first candidate list, a count value m, a count value j, a second candidate list, and a count value n. The count value i represents an entry in the path control table 1321 to be subjected when a retrieval process is performed under the first condition. The first candidate list stores an arrangement of entries in the path control table selected as a transmission candidate under the first condition. The count value m represents the number of entries stored in the first candidate list. The count value j represents an entry in the first

candidate list to be subjected when a selection process is performed under the second condition. The second candidate list stores an arrangement of transmissions interfaces selected as transmission candidates under the second
5 condition. The counter value n represents the number of entries stored to the second candidate list. First, the transmission IP packet 1611 is derived from the IP packet queuing section 1317 (step 400). "1" is set as an initial value to the count value i, j and "0" is set as an initial
10 value to the count value m, n (step 401). Then, the flow goes to the step 402.

[0157] Next, the count values i and x are compared and it is decided whether or not the first candidate retrieval has been completed (step 402). Specifically,
15 when the count values i is equal to the count value x , the first candidate retrieval is completed. When the count values i is not equal to the count value x (NO in step 402), the value stored in the destination IP address of the transmission IP packet 1611 is compared with the value
20 stored in the destination network IP address of the i -th entry in the path control table (step 403). When the comparison result provides a match (YES in step 403), the content of the i -th entry in the path control table is copied to the m -th entry of the first candidate list (step
25 404). Thus, the count value m increments by 1 (step 405).

Moreover, when the count value i increments by 1 (step 406), the flow goes to the next path control table entry process. When the value stored in the destination IP address of the transmission IP packet 1611 does not mach
5 the value stored in the destination network IP address of the i -th entry in the path control table (NO in step 403), the flow goes to the step 406, without any change.

[0158] By repeating the process between the steps 402 and 406, each entry is processed sequentially from the
10 first entry in the path control table. When the process reaches the last entry (YES in step 402), the flow goes to the step 407 to decide whether of not the entry exists in the first candidate list.

[0159] Specifically, the existence or presence of the
15 first candidate is determined by the count value m . When the count value m is "0" (YES in step 407), the process ends because no communications interface exists to transmit the transmission IP packet 1611. When the count value m is not "0" (NO in step 407), the first candidate
20 exists and the flow goes to the step 408. Next, it is decided whether or not the count value m is "1" (step 408). When the count value m is "1" (YES in step 408), the communications interface for transmitting the transmission IP packet 1611 is uniquely decided. For that reason, the
25 communications interface recorded in the first entry of

the first candidate list is captured (step 409) and the transmission IP packet 1611 is transmitted using the captured communications interface (step 422). Thus, the process ends. When the count value m is not "1" (NO in step 408), the flow goes to the step 410 because the communications interface used for transmission of the transmission IP packet 1611 cannot be decided uniquely. Thus, the flow goes to the step 410.

[0160] It is decide whether or not the second candidate retrieval has completed in the step 410. Specifically, when the count value j is equal to the count value m, the second candidate retrieval has completed. When the count value j is not equal to the count value m (NO in step), the communications interface of the j-th entry in the first candidate list is first captured (step 411). Next, the communications interface stored in the transmission source IP address of the transmission IP packet 1611 is compared with the communications interface captured in the step 411 (step 412). When there is a match (YES in step 412), the communications interface captured in the step 411 is copied to the n-th entry in the second candidate list (step 413). Thus, the count value n increments by one (step 414). Moreover, the count value j increments by one (step 415). Thus, the flow goes to the process of the next entry in the first candidate list.

When the interface stored in the transmission source IP address of the transmission IP packet 1611 does not match the communications interface captured in the step 411 (NO in step 412), the flow goes to the step 415 without any change.

[0161] By repeating the process between the steps 410 and 415, each entry is processed sequentially from the first entry in the first candidate list. When the process reaches the last entry (YES in step 410), the flow goes to the step 417 to decide the existence or presence of an entry in the second candidate list.

[0162] Specifically, the existence or presence of the second candidate list is decided by the count value n. The count value (n=) "0" means that the communications interface, which can transmit the transmission IP packet 1611 to the first candidate list, exists but is not in the status uniquely decided. The count value (n=) "1" means that the communications interface, which can transmit the transmission IP packet 1611, is uniquely decided. The count value (n=) "2" or more means that the communications interface, which can transmit the transmission IP packet 1611 to the second candidate list, exists but is not in the status uniquely decided.

[0163] When the count value n is "1" (YES in step 416), the communications interface is captured from the

first entry in the second candidate list (step 417). The transmission IP packet 1611 is transmitted using the captured communications interface (step 421) and thus the process ends. When the count value n is not "1" (NO in
5 step 416), the flow goes to the step 419.

[0164] Next, when the count value n is "0" (YES in step 418), the transmission interface is captured from a given entry in the first candidate list (step 419). The transmission IP packet 1611 is transmitted using the
10 captured transmission interface (step 421) and thus the process ends. When the count value n is not "0" (NO in step 418), the transmission interface is captured from a given entry in the second candidate list (step 420). The transmission IP packet 1611 is transmitted using the
15 captured communications interface (step 421) and thus the process ends.

[0165] In the IP header of the transmission IP packet 1611, the destination IP address stores [200.7.8.9] and the transmission source IP address actually stores an
20 identifier "S-IF#2" (not an IP address). In the path control table 1321, the entries 1321-a, 1321-b, and 1321-c correspond to the first candidate. In the execution of the process in the step 407, the first entry in the first candidate list stores the content of the entry 1321-a in
25 the path control table 1321. The second entry in the first

candidate list stores the content of the entry 1321-b in the path control table 1321. The third entry in the first candidate list stores the content of the entry 1321-c in the path control table 1321. "3" is stored as the count value m. In the execution of the process in the step 416, the second candidate list stores "S-IF#2" only as a transmission interface and stores "1" as the count value. That is, the process of the step 418 is executed based on the decision result by the process in the step 417. In the step 418, "S-IF#2" is captured as the transmission interface. The transmission IF decision/delivery processor 1320 transmits the transmission IP packet 1611, using the communications interface S-IF#2 1310-2 corresponding to the path of a path number 3 selected by the scheduler 313-2.

[0166] As a result, the communications interface S-IF#2 1310-2 transmits the transmission IP packet 1611 and the communications interface R-IF 1500-1 in the reception node 101-2 receives the IP packet 1611 via the carrier network GW 1400-3

[0167] In the above-mentioned procedure, the present embodiment includes plural communications interfaces in which the transmission node 100-2 and the reception node 101-2 do not correspond to each other one to one. In the IP network configuration in which an individual IP address

is allocated to each communications interface, IP packets can be transmitted using the path decided by the scheduler 313-2.

5 [0168] The IP routing process can be performed to IP packets, which do not particularly specify a communications interface but want the IP routing in the conventional art, not specified by the communications interface.

10 [0169] In the transmission node of the above-mentioned embodiment, the queuing section, the scheduler, and the path status monitor are constructed respectively. However, all or part of those elements may be constructed using the CPU that operate under the control program.